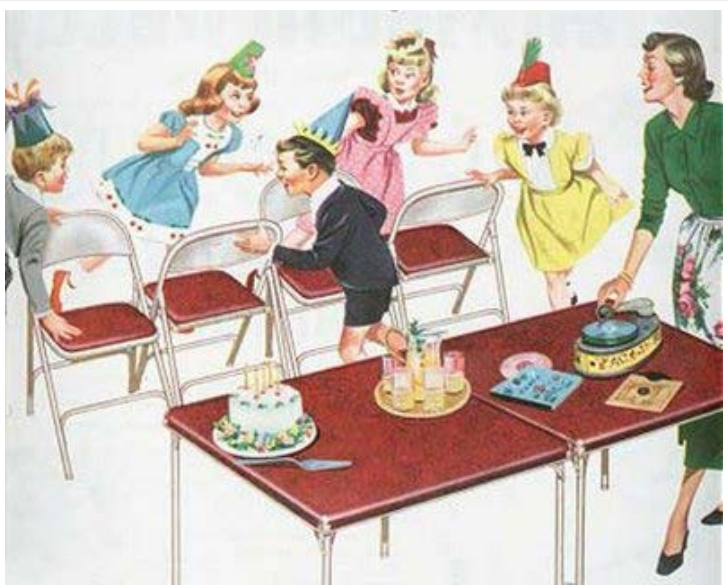
Diese Website verwendet Cookies von Google, um Dienste anzubieten und den Traffic zu analysieren. Deine IP-Adresse und dein User-Agent werden zusammen mit Messwerten zur Leistung und Sicherheit für Google freigegeben. So können Nutzungsstatistiken generiert, Missbrauchsfälle erkannt und behoben und die Qualität des Dienstes gewährleistet werden.

WEITERE INFORMATIONEN OK

An unending variety of topics

Sunday, October 22, 2006

The pigeonhole principle



The pigeonhole principle sounds trivial but is profound. It says that you can't fit **n** pigeons into fewer than **n** holes without cramming at least two pigeons into one of the holes. It says that somebody must lose in a game of musical chairs. In fancier mathematese, which you can safely skip if it sounds like gibberish, the pigeonhole principle says that there is no bijective (or 1-to-1 and onto) mapping (or function) between a set S and any proper subset of S, or for that matter any set smaller than S. [*]

The pigeonhole principle readily proves that there are people in Ohio with the same number of hairs on their head, that you can't eliminate the possibility of hash collisions when the set of possible input data is larger than the possible outputs, that if there are at least two people in a room then there must be at least two people in that room with the same number of cousins in the room, and that a lossless data compression algorithm must always make some files *longer.* This is just the tip of the iceberg of what the pigeonhole principle can help prove.

First let's prove that if there are **n** people in the room, where **n** is at least two, then there must be at least two (but not necessarily **n**) people in that room with the same number of cousins (of the same degree or less) in the room. In mathematese (which you can again ignore if you wish) being a "cousin" a symmetric relationship (if I'm your cousin then you're mine) and non-reflexive (I'm not my own cousin). The proof actually works for any such relationship: twins separated at birth, members of the opposite sex, and so on. Here it is, another proof by contradiction:

(1) Assume that each person in the room has a different number of cousins in the room;

(2) Putting the pigeons in their holes, one person must have no

Pages

- Home

About Me

Nick Szabo

"A premier thinker about history, law and economics, and the lessons they have for security." -- Adam Shostack, Emergent Chaos

"Szabo comes out with these essays that leave me in awe." -- Brian Dunbar

"Reading material that is eclectic, challenging and endlessly fascinating." -- Sean McGrath, Propylon

"Like most blogs worth my attention, this blog is updated only infrequently. That is because the authors of blogs worth my attention only post when they have something to say that is true, relevant and not already known by their audience. Most of the human race does not have the skill to know when an idea has these three properties. The skill is particularly rare in the fields of politics and economics, which is why this blog is such a rare and valuable thing." -- Richard Hollerith

View my complete profile

cousins, and another one cousin, and so on; the **n**th person must have **n-1** cousins. That fills up all the holes with no doubly-stuffed holes, right?

(3) Yes, but if the **n**th person is a cousin with every other person in the room, then there can be nobody in the room with no cousins, so sequence (2) is an impossible contradiction, proving that in fact statement (1) is false: there must be at least two pigeons in the same hole, i.e. at least two people with the same number of cousins in the room.

Next let's prove (again by contradiction) that any lossless compression algorithm must make some file larger:

(1) Since it's a compression algorithm, by definition at least one file will be made shorter. Call the shortest such file **F** of size **m** which compresses to size **n** (i.e. **n**<**m**);

(2) furthermore assume no file will be made longer;

(3) thus every file of size **n** or less is incompressible;

(4) since no files are made longer, thus there must be at least **2^n +** 1 files that compress to size **n** or less: all files of size **n** or less plus file **F;**

(5) but we can't fit **2^n +** 1 pigeons into **2^n** holes without putting two pigeons in one of the holes: in other words, at least two files must compress to the same output file, which is lossy compression;

(6) thus assumption (2) is false: if the compression is lossless some file must be made longer.

Lossless compression nevertheless can be quite effective because data is usually far from uniformly distributed among the **2^n** possibilities. The most common patterns are given the shortest codes, resulting in overall compression in almost all real files that aren't already compressed, encrypted, or otherwise random.

The proofs on hairs and hash collisions are left as exercises for the student. (Don't you hate when math textbooks say that?)

Blog Archive

- ► 2018 (1)
- ► 2017 (3)
- ► 2016 (4)
- ► 2015 (3)
- ► 2014 (3)
- ► 2013 (3)
- ► 2012 (8)
- ► 2011 (12)
- ► 2010 (9)
- ► 2009 (29)
- ► 2008 (55)
- ► 2007 (47)
- ▼ 2006 (130)
  - ► December (4)
  - ► November (7)
  - ▼ October (15)
    - Things as authorities
    - Fifty years ago today
    - The pigeonhole principle
    - Three kinds of jurisdiction
    - Citron's "data reservoirs": putting liability at t...
    - How to succeed or fail on a frontier
    - 600,000 excess deaths in Iraq since U.S. invasion
    - Visual mathematics
    - Prosecute U.S. Justice Department lawyers as war c...
    - Hungarian Revolution anniversary
    - Liberty and habeas corpus (II)
    - New fun legal specialty
    - The Lex Gabina: ancient

An old pigeon farm, back when people actually ate these nasty critters. Presumably all the pigeons fit into all the holes.

[*] Update: this holds for finite sets; the technicalities are a bit different for infinite sets. See comments.

Posted by Nick Szabo at 3:23 PM

Military Commissions Act?

Hello Kitty people

Liberty and habeas corpus (I)

► September (15)

► August (20)

► July (5)

► June (10)

► May (4)

► April (3)

► March (11)

► February (10)

► January (26)

► 2005 (44)

11 comments:

Anonymous said...

You should credit Dirichlet. It's "Be Nice To Dead Mathematicians Month."

10:40 PM

Nick Szabo said...

This is just a blog, not an academic journal, and Johann Peter Gustav Lejeune Dirichlet has been well-credited with the principle elsewhere. Most academic papers that use the principle in proofs don't bother to credit him either.

While we're crediting, the game of musical chairs seems to long predate Dirichlet, being traditional in many cultures.

10:47 AM

Tim Burke said...

I really like the article, but it should be noted that your claim in the first paragraph is only true for finite sets:
the pigeonhole principle says that there is no bijective (or 1-to-1 and onto) mapping (or function) between a set S and any proper subset of S, or for that matter any set smaller than S.

The first part of that sentence breaks down for infinite sets. Consider the set of positive integers(Z+) and the set of even integers (E). E is certainly a proper subset of Z+ and yet the map f(x) = 2*x is a bijection between the two.

The second part is true if you interpret smaller set in the sense of cardinality but then it's trivially true since cardinality can be defined in terms of the possibility of bijections.

2:18 PM

Liron Shapira said...

Good article. One detail:

"In fancier mathematese, which you can safely skip if it sounds like gibberish, the pigeonhole principle says that there is no bijective (or 1-to-1 and onto) mapping (or function) between a set S and any proper subset of S"

The exception is when a set S and a proper subset have the same infinite cardinality.

5:03 PM

Ken Lubin said...

but isn't the "n people in a room" example bullshit?

It's operating with the unstated assumption that all of the cousins of each person in the room is also in the room. Without that assumption, it is false. The nth person could have n^n cousins, and none of them need to be in the room.

12:41 AM

Nick Szabo said...

Ken, you're parsing the sentence wrong. The problem is about "the same number of cousins...in the room" not the total number of cousins a person has.

12:13 PM

Nick Szabo said...

Tim and Liron, thanks for the correction.

12:14 PM

oscabat said...

Excellent article. It's a fascinating principle. Your example reminds me of Wikipedia's "Birthday Paradox." Definitely check it out.

3:05 AM

Anonymous said...

Isn't the cousins proof a bit too complicated? Here's a simpler direct proof without pigeonholes:

Divide the people in the room into groups of mutual cousins. Either any such group contains >= 2 people, then those have the same number of cousins. Otherwise, each group has size 1, i.e. every party guest has zero cousins (and there are at least two guests by assumption).

8:30 AM

Anonymous said...

I take that proof back. It seems like I got the family relations wrong, which is pretty embarassing. (I.e. if A and B are cousins of C, then A is not necessarily the cousin of B. They might be sisters.)

10:01 PM

SSU said...

I teach a liberal arts college math course online. Last week I had students find websites they found interesting/useful about the pigeonhole principle (because that was what they were studying that week). I just wanted to let you know that one student used this post for their webpage (they each had to find a different one).

Good Job!

12:44 PM

Post a Comment

Newer Post Home Older Post

Subscribe to: Post Comments (Atom)

Simple theme. Powered by Blogger.